Chapter 1

Cloud Concepts

MICROSOFT EXAM OBJECTIVES COVERED IN THIS CHAPTER:

DESCRIBE CLOUD SERVICES

✓ **Identify the benefits and considerations of using cloud services**

- Identify the benefits of cloud computing, such as High Availability, Scalability, Elasticity, Agility, and Disaster Recovery
- Identify the differences between Capital Expenditure (CapEx) and Operational Expenditure (OpEx)
- Describe the consumption-based model

DESCRIBE CLOUD SERVICES

✓ **Describe the differences between categories of cloud services**

- Describe the shared responsibility model
- Describe Infrastructure-as-a-Service (IaaS)
- Describe Platform-as-a-Service (PaaS)
- Describe Software-as-a-Service (SaaS)
- Identify a service type based on a use case

DESCRIBE CLOUD OBJECTIVES

✓ **Describe the differences between types of cloud computing**

- Describe cloud computing
- Describe Public cloud
- Describe Private cloud
- Describe Hybrid cloud
- Compare and contrast the three types of cloud computing

The first objective in the Microsoft Azure AZ-900 Certification Exam covers basic cloud concepts. These concepts lay a foundation for understanding why companies choose cloud computing and what types of services are available in Azure. These concepts include the various cloud computing models in Azure, the economic benefits of using Azure, and the three primary service categories in Azure—software-as-a-service (SaaS), infrastructure-as-a-service (IaaS), and platform-as-a-service (PaaS).

First, we'll explore cloud computing.

Understanding Cloud Computing

Microsoft currently offers three cloud computing solutions: Microsoft Azure, Microsoft 365, and Microsoft Dynamics 365. Azure, which is covered on the AZ-900 Certification Exam, provides a broad spectrum of cloud services. These services encompass both server-based and end user–based computing services, along with database services and analytics, artificial intelligence, networking, infrastructure, and much more.

The second Microsoft cloud offering, Microsoft 365, is geared primarily toward providing end-user SaaS solutions like Windows, Office, SharePoint, and OneDrive. Microsoft Dynamics 365 encompasses enterprise resource planning and customer relationship management applications. Microsoft 365 and Microsoft Dynamics 365 are not covered on the AZ-900 exam.

Both Microsoft cloud offerings enable organizations to eliminate computing infrastructure that they might normally host themselves. Larger organizations typically host their own servers, networking equipment, and other IT resources in a *data center*, which is a facility specifically designed and constructed to house servers and other IT hardware and related infrastructure. Some organizations maintain their own data centers, whereas others contract with a third-party data center provider to host their IT equipment and resources.

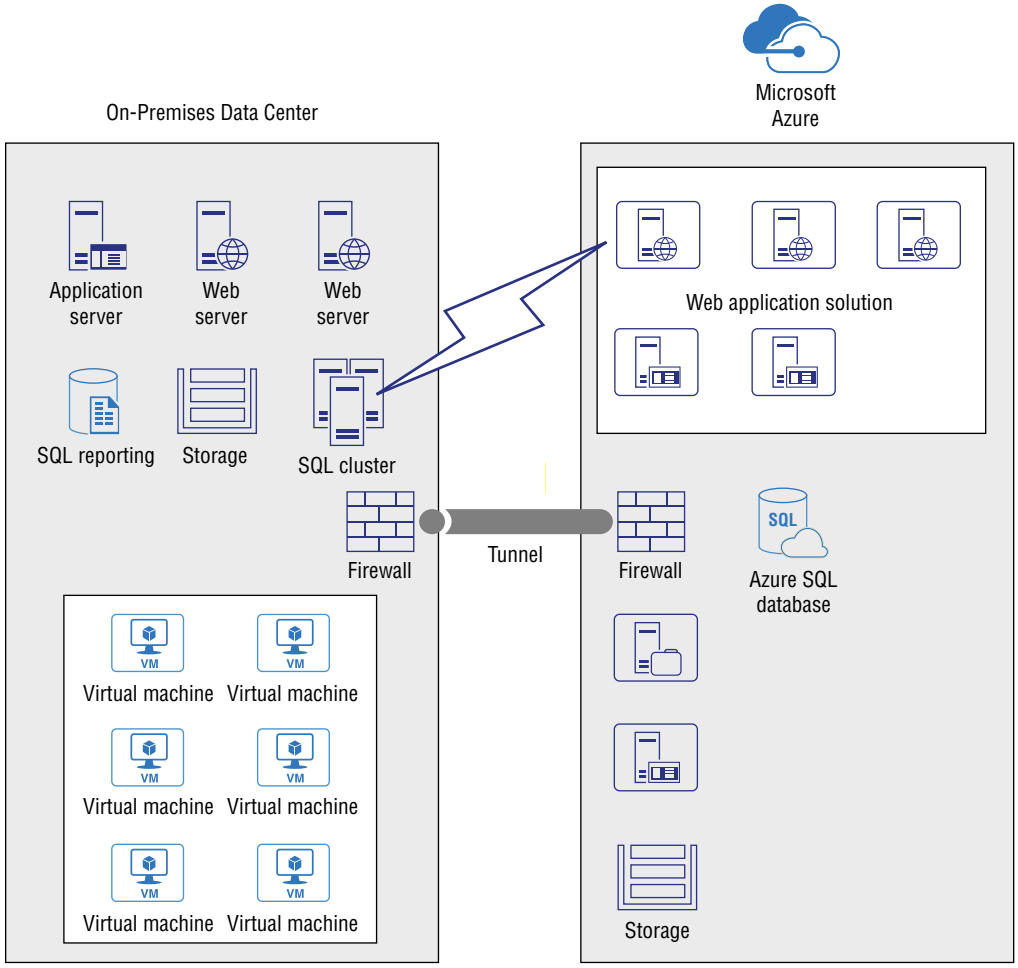
Smaller organizations generally either use a third-party data center or place their servers and other IT infrastructure in one or more *server rooms*, which are essentially very small data centers housed inside the company's facility.

A cloud offering such as Azure enables organizations to move some—if not all—of their servers, networking equipment, and other IT resources into a data center managed by

another company. In the case of Azure, Microsoft owns and maintains numerous data centers around the world to host these resources for all sizes of organizations. Managing these resources then becomes a shared responsibility between the organization and Microsoft. The extent of that shared management depends on the scope of what Microsoft is hosting and the services the organization is using in Azure.

Figure 1.1 shows an example of an organization that is hosting some of its IT infrastructure and services in Azure. As Figure 1.1 illustrates, some of the organization's IT resources remain on site in their own data center, whereas other resources are hosted in Azure, and services interact between the two environments.

FIGURE 1.1 A hybrid cloud scenario



Whatever the case or the extent of services hosted in the cloud, offloading these resources to a cloud provider like Microsoft offers several benefits, which are discussed in the next section.

Benefits of Cloud Computing

Leveraging a cloud computing model offers several benefits, both in financial cost and human resources. This section explores these primary benefits.

Economic Benefits

IT hardware, infrastructure, and related resources can be extremely expensive. In an on-premises model where an organization hosts its own IT infrastructure, whether in its own data center or a third-party data center, the organization bears the cost of the hardware, shipping, support, and related costs. The cost is amortized over several years, sometimes longer than the useful life of the hardware. This type of purchase is a *capital expenditure (CapEx)*, which is money spent by an organization to acquire or maintain fixed assets. Most organizations carefully budget their capital expenditures and require a yearlong budgeting process to set the CapEx budget, and then hold strictly to the budget.

With Azure, Microsoft handles the capital expenditures necessary to maintain and grow the service. An organization using Azure services therefore eliminates those capital expenditures and replaces them with *operational expenditures (OpEx)*, which are monthly expenditures that the organization uses to run its operation. For example, rather than purchasing a license for Microsoft Office for each user (which would be a capital expense), the organization pays a monthly per-user fee for Microsoft 365 (an operational expense). Instead of incurring a relatively large up-front cost for the perpetual license, the organization spreads out the cost on a monthly basis.

The move from a capital to an operational expenditure model can eliminate very large up-front costs to deploy hardware, licenses, support contracts, and other resources. The operational model not only avoids those large up-front expenditures, but also enables the organization to spread the cost throughout the year. It also allows the organization to tie the cost to headcount, so if an individual leaves the organization, the corresponding operational cost also goes away (or is simply reallocated to an incoming resource).

Another economic benefit to cloud computing is *economy of scale*, in which a cloud provider can purchase large amounts of hardware at a discount and pass that discount along to its customers.

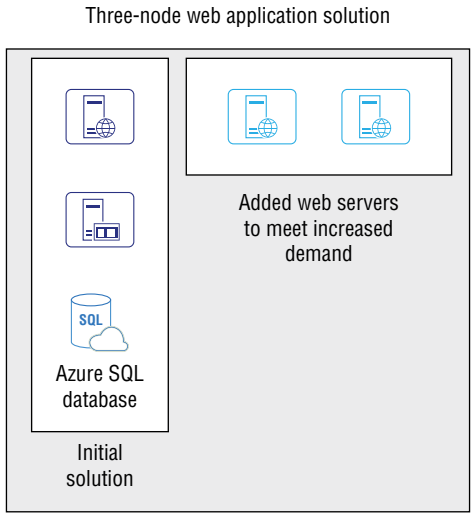
For example, if your organization needed to replace five aging servers, the cost to purchase those servers would be substantial. However, Microsoft purchases servers in large numbers and therefore the cost is less per server and very likely much less than what you would pay for the same equipment. It's not much different from purchasing a case of canned beans at a warehouse store. Per can, a case of 24 is going to be much less expensive than buying a single can. That's economy of scale.

Storage offers another good example of economy of scale. As cloud providers like Microsoft purchase large amounts of storage at a significant discount, that discount is passed to their customers. The result is that storage you purchase from Microsoft is generally much less expensive than storage and related infrastructure that you purchase yourself. As with that can of beans, the cost is shared across many customers, further decreasing the cost to each customer.

Scalability and Elasticity

Two additional benefits of cloud computing are *scalability* and *elasticity*. Scalability is the ability to add computing resources to adjust to increased demand. For example, assume your organization deploys a web farm to handle e-commerce business. For seven or eight months out of the year, your needs are relatively stable. During a peak holiday season, however, you might need several additional servers to handle the increased traffic from sales. You can *scale out* your servers, adding additional ones to meet the increased demand, then *scale in* (eliminate the additional servers) when the peak season is over. This is called horizontal scaling. Instead of incurring the up-front capital expenditure cost of the equipment, you have an increased operational cost only while you are using those servers. Figure 1.2 illustrates an example of horizontal scaling.

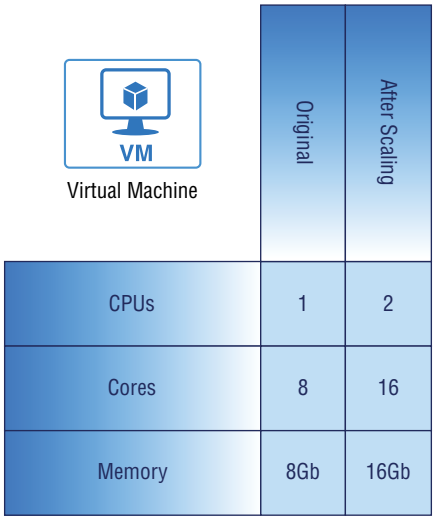
FIGURE 1.2 Horizontal scaling adds additional resources when they are needed and removes them when they are no longer needed.



There are two types of scalability. The previous example described scaling out and its reverse, scaling in. *Scaling up*, also called *vertical scaling*, refers to adjusting capacity in existing resources to accommodate demand changes. For example, increasing the amount

of memory available to a virtual server is an example of scaling up. Adding more processor cores to an existing server is another example of scaling up. As with horizontal scaling, you can go the other way with vertical scaling. When you remove the extra memory when you no longer need it, you're *scaling down*. Figure 1.3 illustrates an example of vertical scaling.

FIGURE 1.3 Vertical scaling adjusts the capacity of existing resources to accommodate demand changes.



Scaling Up a Virtual Machine

Vertical and horizontal scaling can be manual processes, requiring you to take specific action to scale resources. Although it's relatively easy to scale resources in Azure, automatic scaling is often desirable. For example, assume you advertise a product during a television show that airs at 11:00 p.m. This drives consumers to your site, where demand starts to increase around midnight. Your IT staff is all sound asleep. Who is going to notice the demand and take steps to scale accordingly?

Automatic scaling, called *elasticity*, enables Azure to scale resources for you without interaction. Resources can be scaled automatically based on CPU usage, memory usage, storage usage, and so on. Autoscale is the Azure service that enables you to configure automatic scaling. You configure the parameters with Autoscale through rules that you create. When the thresholds defined in the rules are reached, Autoscale handles the process of scaling as defined in the rule.

See `https://docs.microsoft.com/en-us/azure/azure-monitor/platform/autoscale-get-started` for more information on configuring and using Autoscale.

Azure makes it very easy to scale resources to meet demand requirements, and Autoscale—once configured—can make it almost effortless to scale resources. The capability to rapidly adjust resources to meet demand is one aspect of *cloud agility*. Another aspect of cloud agility is the capability to quickly adapt to changing business requirements. The latter is not a concept covered in the Azure Fundamentals certification but is nevertheless an important concept to consider when evaluating cloud computing options.

High Availability

High availability (HA) describes a system that is available for use without significant outages and that is generally backed by a *service level agreement (SLA)*. For example, if a service has an SLA of 99.9 percent, the service is guaranteed to be available 99.9 percent of the time. Translated to the real world, that means the service can be unavailable no more than 43.2 minutes in a 30-day period to meet the 99.9 percent SLA for that month. A financially backed SLA provides a credit for the time in which the service was unavailable.

Service level agreements are discussed in detail in Chapter 6, "Azure Pricing, Service Levels, and Lifecycle."

There are many reasons why a service might become unavailable. Servers might go down, a network issue might prevent traffic to or from the servers, a server application might fail, or a peripheral service might fail. For example, if your line-of-business application relies on a database to host its data and the database goes offline, your application will likely not be available.

It's important to note that "available" does not necessarily translate to performance. If an Azure service is available but at reduced performance, it is nonetheless considered available in the context of the SLA.

Fault Tolerance

The term *fault tolerance* describes a characteristic of a system that enables it to continue functioning when one or more components of the system fails. For example, a typical SharePoint farm consists of at least one database server, a web server, and an application server. These servers together provide the SharePoint services that users consume. If the web server goes down, the service becomes unavailable. To make the SharePoint farm more fault tolerant, you can add a second web server and balance traffic between them. So, if one web server goes down, the other continues to serve web requests. Users might notice a

degradation in responsiveness, but the service remains functional. The SharePoint farm is now fault tolerant to a degree. Figure 1.2 (shown earlier) illustrates this example.

In this SharePoint example, the farm is not fully fault tolerant because points of failure remain in the single application server and database server. When building a fault tolerance strategy for a service, you should consider other points of potential failure, such as a single network path.

Disaster Recovery

Fault tolerance generally applies at the component level of a service. For example, adding a second web server, ensuring that a virtual machine can quickly fail over to another instance, or creating a clustered database instance are examples of fault tolerance. Fault tolerance generally comes into play when a single resource fails.

Disaster recovery refers to the process of recovering from a situation where multiple systems or services fail. For example, assume that your company's primary data center is hit by a tornado, destroying all the IT infrastructure and services hosted at that location. This is certainly a disaster. Recovering from that disaster, however, might be as difficult as setting up all new servers and restoring their configuration and data from backups, or it could be as (relatively) simple as pointing all of your users to a backup data center where all of your infrastructure has been actively duplicated, updated, and ready to become your primary data center.

There is no right answer for a disaster recovery strategy, and it is very much situational and tied to your business continuity needs and defined by the IT services you provide. A very small company, for example, might only need a complete set of backups of its only server and data so that it can quickly restore to a new server. A large organization naturally requires a much more complex disaster recovery plan that can include multiple data centers, active mirroring of services between data centers, and much more. Many organizations are turning to Azure to not only provide a higher level of fault tolerance than they could otherwise implement on-premises, but to implement a disaster recovery environment in Azure for their on-premises systems. Other organizations are turning to Microsoft 365 and Azure to host all of their IT services, with no on-premises IT infrastructure at all, to attain a high level of flexibility, elasticity, fault tolerance, and disaster recovery.

Chapter 2, "Azure Core Services," describes key concepts and services in Azure that provide both fault tolerance and disaster recovery capabilities.

Human Resources

How cloud computing affects human resources is not a major topic on the AZ-900 exam, but the topic does bear some discussion in the context of the benefits of cloud computing.

Organizations move to a cloud model for many reasons: to leverage services that would be very difficult or expensive to deploy on their own, to gain cost savings through economies of scale and shared responsibility, to gain higher elasticity and scalability, or to implement a more effective disaster recovery plan, to name a few.

The indirect consequence of outsourcing IT infrastructure and services to a cloud provider like Microsoft always has some effect on the IT staff who are managing the infrastructure and services that move to the cloud. The following list describes three common scenarios:

- **Staff reduction:** Outsourcing to a cloud provider can potentially reduce the number of IT staff you need to manage those services because Microsoft manages some, if not most, of the infrastructure and services. However, there is a shared responsibility between the organization and Microsoft, and some staff is always required to manage at least some aspects of your cloud footprint.
- **Staff repurposing:** A better approach can be repurposing your staff to more strategic purposes and tasks. Instead of "keeping the lights on," your staff can focus on optimizing the services hosted in the cloud, rolling out new services, and in general playing a more strategic technology role for the business.
- **Staff increase:** It's also possible in some scenarios that an organization will need more IT staff as they transition to the cloud. New services can mean additional staff to direct and manage those services. For example, if you are new to AI and leveraging the AI-related services in Azure, you'll need staff to implement and manage those services.

Financial Models

The section "Economic Benefits" earlier in this chapter briefly described different financial models associated with deploying and managing IT services to meet business needs. The topic bears some additional discussion to further explore the economic benefits offered by cloud computing.

You learned earlier in this chapter that organizations generally budget in two specific areas, capital expenditures and operational expenditures. Organizations typically budget once a year for capital expenditures, which are funds they spend to acquire hard assets like servers and other infrastructure hardware, laptops for employees, company vehicles, buildings and building improvements, and so on. Operational expenses are those funds

the organization spends on a monthly basis to operate the business. Outside the context of cloud computing, operational expenses can include electricity and other utilities, fuel for the company vehicles, and mundane things like snacks for the break room.

Organizations that host their own IT infrastructure and related services almost always incur capital expenditures to deploy those services. The cost of servers, routers, switches, firewalls, software licenses, support contracts, and similar items are usually capital expenditures, and in most cases, very expensive expenditures. Moving to a cloud solution that is paid for via a monthly subscription service based on consumption represents a shift to an operational expense. This shift from capital to operational can mean a huge savings for an organization, as well as greater flexibility and agility.

For example, assume an organization wants to create a proof of concept that requires a dozen servers, storage, server application licenses, a new firewall device, a new load balancer, and other resources. Deploying that as an on-premises solution would require the up-front purchase of those resources. Let's assume that the price for all of that is $500,000.

Deploying that same proof of concept in Azure could be a small fraction of the cost of the on-premises solution. Why? Instead of purchasing servers, you create virtual machines in Azure, which you only pay for while they are turned on. If you purchase the hardware yourself, you incur the cost whether or not those servers are ever operational.

The other infrastructure, including firewalls, load balancers, and other resources, are also paid for on a consumption basis. So, you stand up the proof of concept, evaluate and determine your go-forward strategy, and then turn off all those services when the project evaluation is completed. You've perhaps paid a few thousand dollars for a month of consumption, rather than the $500,000 investment required by the on-premises solution.

The previous example illustrates both the agility and cost savings associated with moving to a cloud platform. It also highlights the benefits of moving from a capital expenditure model to an operational expenditure model. You avoid large, up-front capital expenses and instead transition to monthly operational costs based on your consumption of Azure services.

Cloud Computing Models and Responsibilities

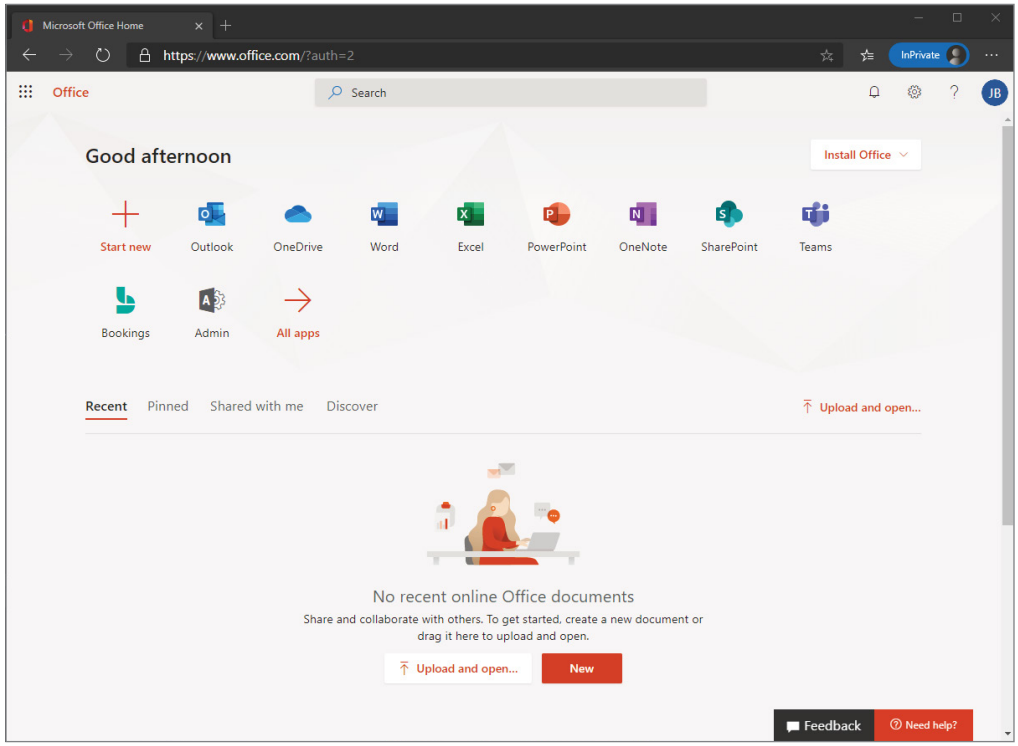
Azure provides three cloud computing models, and each of these results in different levels of shared responsibility between Microsoft and Azure subscribers. The following sections explore these models and how each varies in shared responsibility.

Software-as-a-Service

Azure services fall generally into three categories. The first of these categories is *software-as-a-service (SaaS)*.

SaaS is a licensing model in which an organization licenses software on a subscription basis from a provider that centrally manages and distributes the software. A perfect example of SaaS is Microsoft 365, where Microsoft provides Office applications (Word, Excel, and others) on a subscription basis (see Figure 1.4). Rather than purchase a perpetual license for Microsoft Office for each user, an organization pays a monthly subscription fee per user to enable that user to access and use the software.

FIGURE 1.4 Microsoft 365 is an example of an SaaS offering.



With a perpetual license, the organization is responsible for deploying the software to the user's device and managing updates. There are exceptions, but generally the organization must purchase a new license to upgrade to the next version of the software.

With an SaaS subscription, the user often uses the software through a web interface. With a Microsoft 365 subscription, for example, the user can use Word, Excel, and the other Office applications through a web browser. A Microsoft 365 subscription gives the user the capability to install the applications on a device and use it locally as well.

A typical Azure SaaS offering is Azure SQL Database. Rather than deploy a physical server or virtual machine, install Microsoft SQL Server on that server, and create a database, you simply create a SQL database in Azure. You don't need to worry about managing the server or the SQL Server application—you only manage the database itself.

SaaS offers a number of benefits:

- **Operational expense:** Instead of incurring a capital expenditure to obtain a perpetual license, you move licensing cost into an operational expenditure model. This not only reduces potentially high, lump-sum payments but also enables you to more easily tie the cost of providing IT services to active users.
- **Updates:** With most perpetual licensing models, you must purchase upgrade licenses or even another full license when you want to move to the next version of an application. With an SaaS offering, the provider generally updates the application periodically as a part of the service offering, giving your users access to new features as part of their existing subscription.
- **Deployment and manageability:** You do not need to deploy or manage software with an SaaS offering. The provider manages the application and you simply consume it. Some shared responsibility still exists, however, in that you must manage access to the software through how you apply licensing. However, the responsibility of managing the software itself and the infrastructure on which it runs falls to the provider.

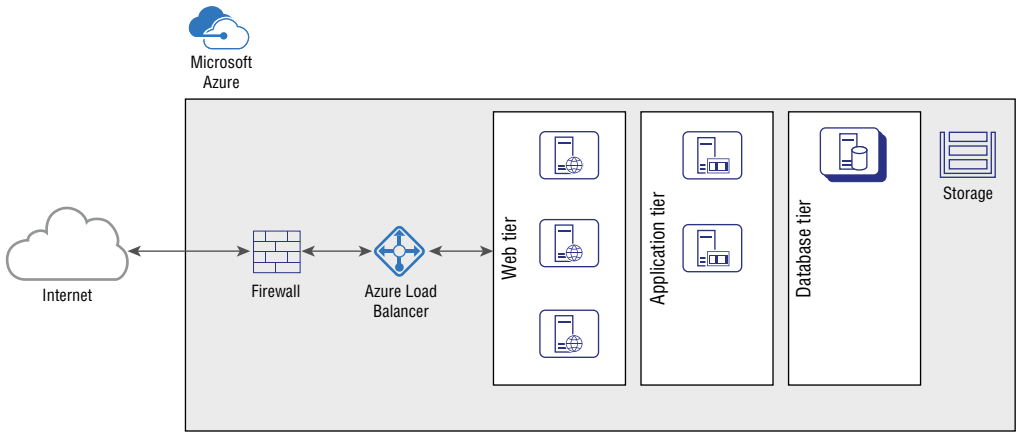
Infrastructure-as-a-Service

The second service category in Azure is infrastructure-as-a-service (IaaS). As its name implies, IaaS enables you to deploy and use infrastructure components in Azure. A primary example of IaaS is *virtualization*, in which you deploy virtual servers in Azure rather than physical servers in your own data center.

For example, assume you want to stand up a SharePoint farm consisting of three web front ends, two application servers, and a two-node SQL database cluster. You have to set up your own farm because you need customization and other capabilities not available in the SharePoint features included with your Microsoft 365 implementation. So, instead of purchasing the seven servers needed for the farm, licensing the operating systems and application software, buying storage, and setting up the farm in your own data center, you stand up seven virtual machines in Azure (see Figure 1.5). These virtual machines are provisioned and run on virtualization hardware that Microsoft manages. Microsoft also manages the networking infrastructure, physical storage, and other resources needed by the SharePoint solution.

The first advantage of this solution is that you don't incur a capital expenditure for the hardware or for the operating system licenses. You also only incur operational costs when the servers are running. So, if you spend a week getting the virtual servers and the farm configured, then shut them down for a couple of weeks while you work on other aspects of the project, you don't incur any costs for the time those servers are turned off.

FIGURE 1.5 A simplified illustration of a virtualized SharePoint farm in Azure



Scalability is another important aspect of IaaS. Using the same SharePoint example, you can easily add resources to the web front ends (vertical scaling) or add additional web front ends (horizontal scaling) to accommodate increased traffic to the farm.

Flexibility is another important benefit of IaaS over using dedicated hardware. Assume your team spends a few days standing up the virtual machines and getting the SharePoint farm operational, and then the business makes the decision to take the project in a different direction or to cancel the project altogether. You shut down the servers and related services, ending the infrastructure costs associated with the project. You aren't stuck with hardware and other infrastructure that you now need to repurpose for something else.

IaaS generally has a higher shared responsibility than SaaS. Although Microsoft manages the hardware that supports your virtualized infrastructure, you still need to manage the virtual servers, operating systems, and applications installed on the virtual machines. You will learn more about the different responsibilities later in this chapter in the section "Shared Responsibility."

Platform-as-a-Service

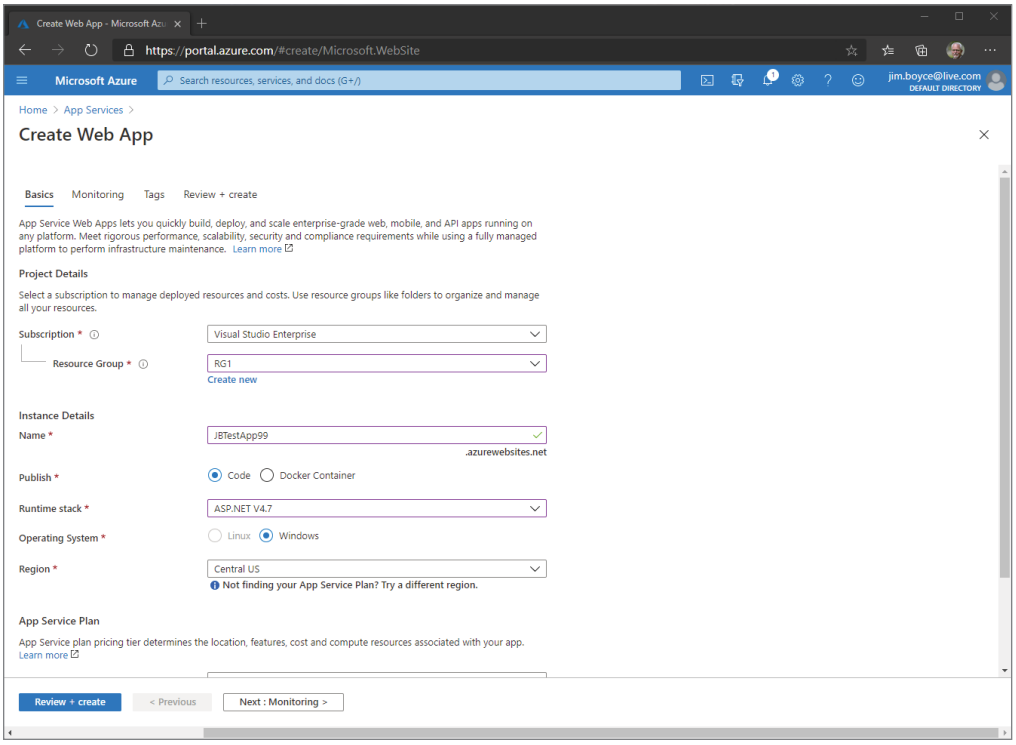
Conceptually, platform-as-a-service (PaaS) is a combination of IaaS and SaaS in that it incorporates both infrastructure and software. Microsoft provides and manages the virtualized infrastructure (virtual machines, networking, and so forth) and provides additional software and resources to facilitate application development. So, think of PaaS as providing the capability to quickly develop and deploy web-based applications without the need to manage the underlying servers, operating systems, and other resources that are part of the developed application.

Previously in this chapter I identified Azure SQL Database as an SaaS offering. That is correct, but Azure SQL Database can also be a component of PaaS in that your application might be leveraging Azure SQL Database as part of the solution. The next section, "Shared Responsibility," provides additional details.

For example, assume that your organization needs to develop and deploy a web application that uses a database to store data used by the application, and you want to use ASP.NET for the application. You could deploy some virtual machines, a SQL server to host the database (or use Azure SQL Database to host it), and other required software to provide for ASP.NET integration, and then develop and deploy the web application to those servers. However, you want to streamline your development effort and don't want to have to manage the virtual machines or other resources. Instead, you want to focus specifically on the web application.

In this scenario, you turn to Azure App Services to provide the underlying virtual machines, operating systems, software, and other resources. Figure 1.6 shows an example of a web app service being created in Azure. As the figure illustrates, this web app is called JBTestApp99 and uses ASP.NET 4.7 on Windows.

FIGURE 1.6 Creating a web app named JBTestApp99 to use ASP.NET 4.7 and run on Windows



Even though this solution is running on Windows and uses ASP.NET, you don't need to deploy or manage the Windows environment, ASP.NET, or other resources. Instead, you focus solely on the application development.

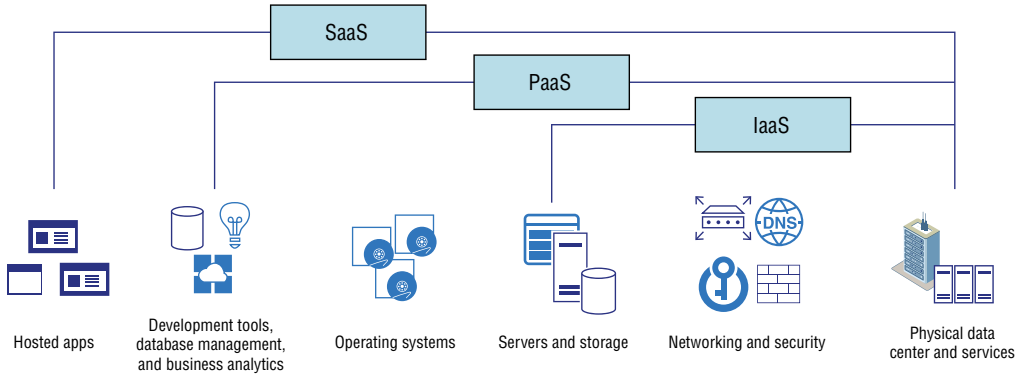
There are many other aspects to creating web applications using Azure, including the capability to create multiple app service plans, configure automatic monitoring and application insights for the app, and more. For the purposes of the AZ-900 exam, however, just understand that PaaS provides a development platform that you can use to create and deploy many different types of web applications without worrying about deploying or managing the servers and other infrastructure that support that application.

Shared Responsibility

In an on-premises IT scenario, your organization is generally responsible for all aspects of a service, including physical space, hardware, power, cooling, operating systems, applications, networking, and more. With cloud computing, you have a shared responsibility with the cloud provider. In the case of Azure, Microsoft manages certain aspects of the services that you host in Azure, and you are responsible for other aspects.

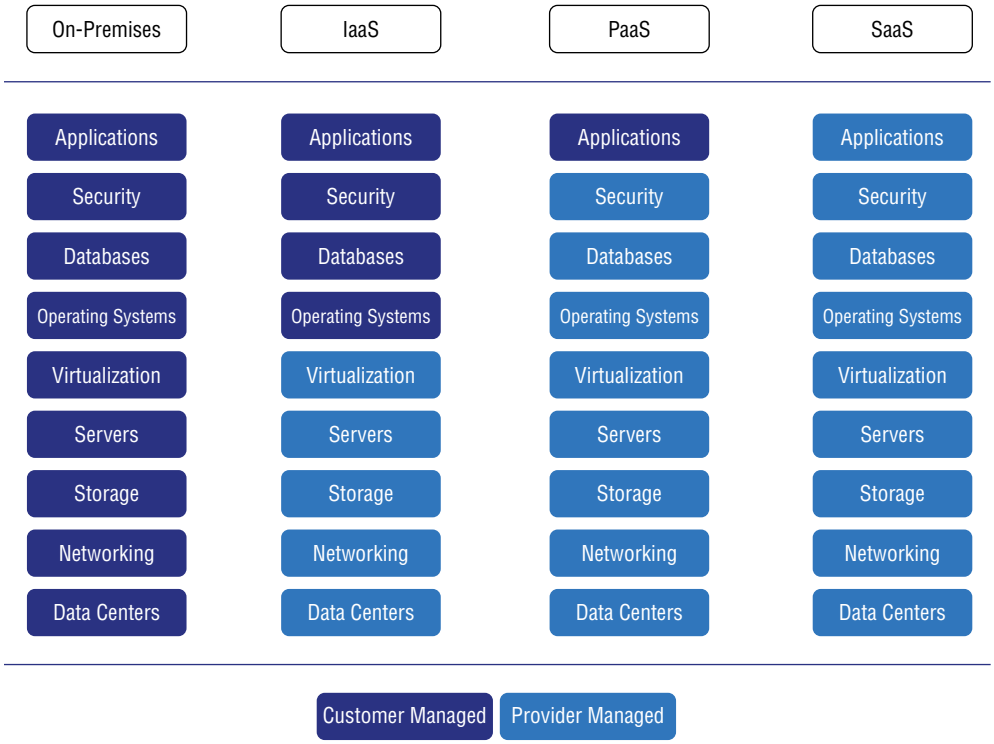
Figure 1.7 shows the relationship of different cloud computing elements and how they fit into the categories described previously in this chapter. As the figure illustrates, different categories overlap, with IaaS being a subcomponent of both PaaS and SaaS.

FIGURE 1.7 IaaS, PaaS, and SaaS are all categories of cloud computing and share a nested relationship.



Each category of cloud computing in Azure involves different levels of responsibility from your organization and Microsoft. Figure 1.8 illustrates the various levels of responsibility with each of these categories, with IaaS having the highest level of responsibility for the customer and SaaS having the least responsibility.

FIGURE 1.8 The three categories of Azure cloud computing have different levels of shared responsibility between your organization and Microsoft compared to on-premises/self-hosted.



Public, Private, and Hybrid Cloud Models

A key topic covered on the AZ-900 exam is the difference between cloud computing models. This section explores the three models: public, private, and hybrid.

Public Cloud

A public cloud is one in which the services offered in the cloud are available for public use over the Internet. This doesn't mean that the services you deploy in Azure are available to the public. In a public cloud, many boundary layers segregate your services from those hosted in Azure by other organizations. The services and data that you host in Azure are secure and inaccessible by people outside of your organization (unless you specifically provide guest access to certain services, such as your company website).

Using a public cloud to host some of your IT services doesn't mean that all of your services are moved to the cloud. You might put some of your services in Azure but still maintain other services on-premises in your own data center.

Private Cloud

A private cloud is one in which the cloud serves a single organization, whether hosted in your own data center or by someone else. A private cloud offers many of the same benefits of scalability, elasticity, and other aspects of a public cloud. Because a private cloud is dedicated to one organization, it offers some additional capabilities to meet regulatory requirements because you can impose controls and processes in the private cloud that would not necessarily be available in a public cloud. Security is also another aspect where a private cloud offers some advantages. Many organizations have regulatory requirements to prevent certain types of data from being exposed to the Internet, and hosting that data in a private cloud can help that organization meet that requirement.

Although a private cloud offers many benefits, it also has several potential drawbacks. For example, hosting services in a private cloud is often no less expensive than simply hosting those same services on-premises in your own data center because the hardware and most of the infrastructure is dedicated to your organization, even if it is hosted by a third party. You either pay for the hardware yourself or the hosting provider passes the cost on to your organization in some fashion. Any cost savings you realize from a third-party private cloud generally come from the fact that the hosting company manages the data center for multiple customers, which provides some economy of scale.

Hybrid Cloud

A hybrid cloud is one in which an organization integrates on-premises services with cloud-hosted services, whether in a public or a private cloud. Using Microsoft 365 as an example, you might continue to host your own Exchange servers in your data center to provide certain internal functions but also have mailboxes hosted in Exchange Online. The on-premises Exchange servers and Exchange Online service work together to provide a consolidated messaging solution.

An example of an Azure hybrid scenario is one in which you host data on-premises in a SQL Server cluster but have applications hosted in Azure that use that data. The two services work together to provide the end solution to your users.

The key to differentiating between a hybrid cloud scenario and a mix of on-premises and cloud services is whether there is service interaction between your on-premises service and the related service(s) in the cloud. For example, assume you use Microsoft 365 to provide messaging and Office applications to your users and use on-premises servers and applications to provide line-of-business applications such as accounting and billing. If the two environments exist separately with no interaction between services in each environment, then you don't have a hybrid cloud scenario. If you do have interaction between services, then you do have a hybrid cloud scenario.

Summary

The Cloud Concepts objective of the AZ-900 exam is intended to help you understand what cloud computing is, its benefits, and different cloud models without diving into specific Azure services, pricing models, or other details (except where needed peripherally to understand a basic concept).

This chapter covered the following concepts:

- **Benefits of cloud computing:** Cloud computing offers a number of economic benefits, including economies of scale and the opportunity to move from a capital expenditure model to an operational expenditure model. Scalability and elasticity enable you to scale your Azure services to meet changes in demand, adding resources when needed to meet increased demand and scaling back as demand decreases, with associated changes in cost. Cloud computing also provides improved capabilities for both high availability and fault tolerance, enabling your services to respond to issues and outages.
- **Cloud computing models and responsibilities**: The three cloud models—software-as-a-service, infrastructure-as-a-service, and platform-as-a-service—each provide different capabilities with different levels of shared responsibility aligned to each. Trade-offs exist within each model between flexibility and control. For example, IaaS gives you the most control over the services hosted in Azure, whereas SaaS is generally the easiest to manage.
- **Public, private, and hybrid cloud models:** These three cloud models each offer certain benefits and can be used individually or together to achieve business goals through IT services. All of the models offer significant potential savings in both cost and effort over on-premises hosting models. A public cloud generally offers the lowest cost, a private cloud a higher cost but greater flexibility to meet certain regulatory requirements, and a hybrid cloud the capability for integrating systems on-premises and in the cloud.

Exam Essentials

Describe terms such as high availability, scalability, elasticity, agility, fault tolerance, and disaster recovery. High availability is a characteristic of Azure that refers to the availability of specific services, measured as a percentage (available/total time). For example, a service with a service level agreement (SLA) of 99.9 percent is guaranteed to be unavailable no more than 43.2 minutes in a 30-day period. Reduced performance is not counted as unavailable.

Scalability is a characteristic of Azure that enables services to adjust to changes in demand. Vertical scaling refers to adding resources to an existing system, such as adding more memory or CPU cores to a virtual machine. Horizontal scaling refers to adding additional systems, such as adding additional virtual machines. Scaling down and scaling in are the reverse of scaling up and scaling out. Scalability is a function of elasticity.

Agility refers to the capability to quickly deploy services with reduced effort and cost.

Fault tolerance is a characteristic of a system that enables it to tolerate the failure of one or more of its components.

Disaster recovery is the process of recovering systems and data following a major failure or disaster.

Describe the principles of economies of scale. Technical economies of scale are achieved when a cloud provider can purchase a large amount of hardware and other infrastructure at a discount, resulting in a cost savings to its consumers, and/or when those resources are used by multiple organizations, effectively reducing the cost per user or per organization.

Describe the difference between capital expenditure (CapEx) and operational expenditure (OpEx). A capital expenditure (CapEx) is the acquisition of hard assets, such as the purchase of servers and other IT hardware. Operational expenditures (OpEx) are expenses incurred in operating a business, such as monthly consumption fees for using Azure services.

Describe the consumption-based model. In a consumption-based model, an organization pays for the resources the organization and its users consume, generally resulting in a cost savings because costs are incurred only when the resources are in use.

Describe infrastructure-as-a-service (IaaS). IaaS refers to compute, networking, and related services that your organization consumes from a pool of resources hosted by a cloud provider. The provider manages the physical hardware and supporting infrastructure, and your organization manages the operating system, applications, networking configuration, and related services.

Describe platform-as-a-service (PaaS). PaaS describes a system that enables organizations to quickly develop and deploy applications without the need to obtain or manage the underlying hardware or other infrastructure required by the application.

Describe software-as-a-service (SaaS). SaaS describes a model in which an organization consumes software hosted by a cloud provider. The cloud provider manages the application and updates, and the consumer manages access to the applications by its users. Microsoft 365 is an example of SaaS.

Compare and contrast the three different service types. IaaS is generally tied to virtualization and the capability to quickly stand up and manage virtual servers in the cloud with consumer control over the operating system and applications running on the VMs. PaaS abstracts the hardware and underlying support applications (*middleware*) and instead focuses on the interaction between the consumer and the service, simplifying the capability to consume development-related services. SaaS fully abstracts all hardware and application support, enabling the organization's users to use an application without needing to manage any aspect of the application or its underlying infrastructure. Office applications in Microsoft 365 are an example of applications served through an SaaS model.

Describe a public cloud. A public cloud is one in which services are provided to multiple organizations through a publicly accessible network such as the Internet. In general, physical compute and networking hardware, along with other supporting infrastructure, is

shared among the organizations consuming services from the public cloud. Various physical and virtual boundaries securely separate one organization's services and data from those of others. Reduced cost is the primary advantage to a public cloud model.

Describe a private cloud. A private cloud is one in which services are provided to a single organization, whether managed internally by the organization or by a third party. Greater control over systems, applications, and data is the primary advantage of a private cloud.

Describe a hybrid cloud. A hybrid cloud model is one in which non-cloud services hosted on-premises directly interact with services hosted in either a public or a private cloud.

Review Questions

1. Is the underlined portion of the following statement true, or does it need to be replaced with one of the other fragments that appear below?

   As a cloud service, Microsoft Azure enables your organization to budget IT infrastructure costs as a capital expenditure.

   A. as an operational expenditure.

   B. on an annual basis.

   C. using the Azure Pricing Estimator.

   D. No change is needed.

2. The cost per subscriber decreases as the number of Azure subscribers increases. Which benefit of cloud computing does this statement describe?

   A. Agility

   B. Scalability

   C. Economy of scale

   D. Elasticity

3. You are an IT manager for a small company that hosts a web application for e-commerce. The web application uses two web servers and a small database cluster. As demand increases through a peak season, you want to add additional web servers to handle the increased demand, then remove those additional servers as demand decreases. You propose moving the application to Azure. Which of the following statements is true?

   A. The web servers can be moved to Azure but the database cluster must remain on-premises, which represents a hybrid cloud model.

   B. Horizontal scaling enables you to add and remove web servers to meet demand changes.

   C. Vertical scaling enables you to add and remove web servers to meet demand changes.

   D. Azure will automatically add the web servers for you as demand approaches a threshold set by the Azure App Service.

4. You are an IT director for Contoso and are preparing a proposal to your CIO to move all IT infrastructure to Azure. Which of the following is an advantage to moving your infrastructure to a public cloud provider?

   A. You will have complete control over all infrastructure, network, applications, and all other resources in the cloud.

   B. You can scale your infrastructure horizontally or vertically without capital expenditure costs.

   C. You will reduce your operational expenditures.

   D. The cloud provider will manage all infrastructure for you, enabling Contoso to reduce IT staff.

5. Contoso is building a web application that uses a SQL database to store data. Which of the following represents a hybrid cloud scenario? (Choose all that apply.)
   - **A.** A virtual machine in Azure that hosts the web application and a second virtual machine in Azure running SQL Server
   - **B.** A two-node SQL cluster in a third-party data center that hosts the data and two virtual machines in Azure running the web application
   - **C.** A web application hosted in Azure that stores its data in an Azure SQL database
   - **D.** A web application hosted in Azure that sends data to and from a database hosted in Contoso's on-premises data center

6. Is the underlined portion of the following statement true, or does it need to be replaced with one of the other fragments that appear below?

   Moving servers from an on-premises data center to virtual machines in Azure enables you to reduce IT staffing because Microsoft manages the infrastructure for you.
   - **A.** enables you to pay for only the Azure resources you consume on a monthly basis.
   - **B.** represents a platform-as-a-service (PaaS) solution.
   - **C.** enables you to use additional firewall services only available in Azure to protect against security risks.
   - **D.** No change is needed.

7. You are the application development director for Contoso. Your team needs to develop and bring a new web application online quickly with minimal expense. You consider using Azure Web Apps, Azure Functions, and Azure Database for MySQL. Which Azure service category does this represent?
   - **A.** Infrastructure-as-a-service (IaaS)
   - **B.** Platform-as-a-service (PaaS)
   - **C.** Software-as-a-service (SaaS)
   - **D.** Development-as-a-service (DaaS)

8. Your company currently installs Microsoft Office on each user's computer using perpetual licenses that you have purchased from a licensing vendor. You propose to the CIO to transition your users to Microsoft 365 to use Office applications hosted by Microsoft, which enables your organization to work with documents in a web browser and also receive Office application updates automatically. This represents which service category?
   - **A.** Infrastructure-as-a-service (IaaS)
   - **B.** Platform-as-a-service (PaaS)
   - **C.** Software-as-a-service (SaaS)
   - **D.** None of the above

9. You are a server administrator for Wingtip Toys, a small company that makes and distributes wooden toys. You manage a custom line-of-business (LOB) application for order management and shipping. The solution is hosted on aging servers in a server room in your

manufacturing facility. You want to eventually move the application's functions to Microsoft Dynamics 365 and eliminate the custom application. You propose to your manager that you first migrate the servers into virtual machines hosted in Azure to avoid purchasing new, up-to-date hardware. The current application will run on these new VMs. This proposal represents which of the following?

A. Software-as-a-service (SaaS)

B. Infrastructure-as-a-service (IaaS)

C. Platform-as-a-service (PaaS)

D. A hybrid cloud scenario

10. You are the CIO for a company and are concerned about the security of your data in the cloud. You need to implement a cloud solution in which you gain the flexibility and agility of a cloud solution but maintain full control of your data and infrastructure. You propose to your CEO that you contract with a third-party cloud vendor to host your IT services, and the infrastructure on which your services will be hosted will not be used by any other organization. This represents which type of cloud model?

A. Public cloud

B. Private cloud

C. Hybrid cloud

D. Both A and B

11. You deploy a custom data analytics application to Azure that includes a single web front end through which the users access the application. At peak times during the day, the web server experiences very high memory usage and temporarily enters an unresponsive state due to a bug in your application. As a stopgap measure while your developers research the issue, you add a second web server and balance the load between the two web servers. Although the service sometimes slows down, the servers are able to independently recover from the memory issue and the service remains available. Scaling out the web servers resulted in improvements in which two areas?

A. Disaster recovery

B. Agility

C. Fault tolerance

D. High availability

12. You deploy a web app using Azure App Services and configure autoscaling for it so that it can request additional compute resources when the app experiences high increases in demand. What is this an example of?

A. Elasticity

B. PaaS

C. Fault tolerance

D. High availability